ELSEVIER

Journal of Clinical Epidemiology 60 (2007) 981–989

Journal of Clinical Epidemiology

REVIEW ARTICLES

A review identifies and classifies reasons for ordering diagnostic tests

Penny Whiting[a], Merran Toerien[a,*], Isabel de Salis[a], Jonathan A.C. Sterne[a], Paul Dieppe[a], Matthias Egger[b], Tom Fahey[c]

[a]*MRC Health Services Research Collaboration, Department of Social Medicine, Canynge Hall, Whiteladies Road, Bristol BS8 2PR, UK*
[b]*Department of Social and Preventive Medicine, University of Bern, Finkenhubelweg 11, CH-3012 Bern, Switzerland*
[c]*Division of Community Health Sciences, University of Dundee, Mackenzie Building, Kirsty Semple Way, Dundee DD2 4BF, UK*

Accepted 19 January 2007

Abstract

Objective: To consider the reasons and context for test ordering by doctors when faced with an undiagnosed complaint in primary or secondary care.

Study Design and Setting: We reviewed any study of any design that discussed factors that may affect a doctor's decision to order a test. Articles were located through searches of electronic databases, authors' files on diagnostic methodology, and reference lists of relevant studies. We extracted data on: study design, type of analysis, setting, topic area, and any factors reported to influence test ordering.

Results: We included 37 studies. We carried out a thematic analysis to synthesize data. Five key groupings arose from this process: diagnostic factors, therapeutic and prognostic factors, patient-related factors, doctor-related factors, and policy and organization-related factors. To illustrate how the various factors identified may influence test ordering we considered the symptom low back pain and the diagnosis multiple sclerosis as examples.

Conclusions: A wide variety of factors influence a doctor's decision to order a test. These are integral to understanding diagnosis in clinical practice. Traditional diagnostic accuracy studies should be supplemented with research into the broader context in which doctors perform their work.

Keywords: Diagnostic test; Review; Low back pain; Multiple sclerosis; Test ordering

1. Introduction

Diagnostic test assessment focuses on determining test accuracy, defined as the ability of a test to correctly distinguish between those with and without the target condition. However, a wide range of factors, which are ignored in traditional test evaluations, influence test ordering. The importance of taking these other factors into account in test accuracy studies has been noted previously [1]. Despite this, we are unaware of any review of reasons for ordering tests. In this article we review broader reasons for test ordering in patients with an undiagnosed complaint in primary or secondary care.

2. Methods

We searched MEDLINE, Embase, PsycINFO, and the International Bibliography of the Social Sciences (full details of the search strategies available from the authors on request) to identify qualitative studies in the area of diagnosis. Our initial searches identified 1,971 studies. We screened these for relevance and categorized relevant articles thematically. A key theme to emerge was test ordering and we chose to focus further on this topic. We supplemented our initial searches by reviewing our extensive personal files. In addition, we identified further studies by screening reference lists of articles already obtained. Although no formal language restrictions were applied, all the articles included in this review are English language. Studies were ordered and reviewed sequentially until saturation was reached, that is, until no new reasons for test ordering were identified.

We applied very broad inclusion criteria for the review: any study of any design that discussed factors that may affect a doctor's decision to order a test in patients with an undiagnosed complaint in primary or secondary care was eligible for inclusion. This included discussion articles as well as primary studies and systematic reviews because we wanted to provide a comprehensive overview of all possible factors that may influence test ordering. Although tests are also done in patients without symptoms or with

This work was supported by the MRC Health Services Research Collaboration.

* Corresponding author. Tel.: +44-117-928-7367; fax: +44-117-331-3914.

E-mail address: merran.toerien@bristol.ac.uk (M. Toerien).

0895-4356/07/$ – see front matter
doi: 10.1016/j.jclinepi.2007.01.012

Box 1 Why do we order diagnostic tests?

1. Diagnostic factors
   - Modify pretest probability of disease
   - Rule in or rule out disease
   - Primary care: rule out or referral to secondary care
   - Secondary care: reach a definite diagnosis
2. Therapeutic and prognostic factors
   - Decide on appropriate treatment
   - Predict subsequent clinical course and assess prognosis
   - Monitor the effects of interventions
3. Patient-related factors
   - Patient preference, e.g., to order testing, not to undergo a test, or for a particular test
   - Patient acceptability and side effects of test
   - Impact of diagnosis (or lack of diagnosis)
   - Consequences of inaccurate test results
   - Patient reassurance
   - Patient demographics
4. Doctor-related factors
   - Clinical experience and confidence in clinical judgment
   - Knowledge regarding test properties
   - Cognitive biases
   - Involvement in research
   - Attitudes to risk taking/fear of uncertainty/reassurance
   - Fear of litigation – defensive testing
   - Response to patients' requests for inappropriate testing
   - Feedback on doctor's test ordering rates compared to others
   - Doctor specialty
   - Working full time vs. part time
   - Time constraints
   - Doctor demographics
   - Professional pride
5. Policy and organization-related factors
   - Primary care practice size
   - Test availability
   - Method of doctor payment
   - Policy and clinical guidelines
   - Use of structured test ordering form
   - Referral process

confirmed disease (e.g., screening, disease monitoring, judging severity of illness, or determining treatment response), we restricted our review to patients with an undiagnosed complaint to focus specifically on the process of symptomatic diagnosis. Inclusion was assessed by one reviewer and checked by a second.

We extracted data from the individual studies on study design (discussion article, primary study or systematic review); type of analysis (qualitative or quantitative); setting (primary care, secondary care, both or general); topic area; and any factors reported to influence test ordering. Data extraction was carried out by one reviewer and checked by a second. We carried out a thematic analysis to synthesize data [2]. This involved identifying all reasons for test ordering described in included studies and producing a spider diagram to group and link related reasons. Five key groupings arose from this process. We then categorized all the identified reasons for test ordering according to these groupings.

To provide a cohesive illustration of how the various factors identified may influence test ordering, we considered the symptom low back pain (LBP) and the diagnosis multiple sclerosis (MS) as examples. We selected these because they involve different diagnostic challenges. For potential MS, the goal is to rule in a serious condition while avoiding false positive diagnoses; whereas for LBP the goal is to identify the minority of patients with organic disease for which there is a treatment that improves a patient's prognosis, while minimizing testing of the majority whose condition has a nonspecific cause. To identify illustrative examples in these two topic areas we carried out Internet searches using Google Scholar (http://www.scholar.google.com) using key terms related to "test ordering," "(low) back pain" and "multiple sclerosis," and searched our own files.

3. Results

We identified 38 studies that discussed factors that may influence test ordering [1,3–38]. Details of these studies are provided in Table 1. The included studies covered a variety of different designs: primary studies ($n = 22$), discussion articles ($n = 14$), and systematic reviews ($n = 2$). Most of the primary studies were quantitative in design ($n = 17$), four were qualitative, and one presented both a qualitative and quantitative analysis. Most studies covered issues of test ordering from a general perspective, although five studies focused on laboratory studies, three on LBP, one on MS, and one on radiology.

Five key factors emerged as reasons for test ordering: diagnostic factors, therapeutic and prognostic factors, patient-related factors, doctor-related factors, and policy and organization-related factors. These factors are summarized in Box 1. Below, we discuss these in more detail, and show

Table 1
Study details

Study details	General features[a]				Factors that may influence test ordering																																	
					Diagnostic							Therapy		Patient								Clinician												Policy and organization						
	Study design	Analysis	Setting	Topic area	Modify pretest probability	Rule in or rule out disease	1° Care: rule out or referral	2° Care: reach a definite diagnosis	Watchful waiting	Decide on appropriate treatment	Predict course and assess prognosis	Patient preference	Patient acceptability	Impact of diagnosis	Patient reassurance	Patient demographics	Consequences of test results	Clinical experience	Knowledge of test properties	Cognitive biases	Involvement in research	Risk taking/reassurance	Fear of litigation	Feedback on test ordering	Response to patients' requests	Working full time vs. part time	Clinician Specialty	Time pressures	Clinician demographics	Professional Pride	GP practice size	Test availability	Method of clinician payment	Policy and clinical guidelines	Use of structured test form	Referral policy
Asch et al. (1990) [31]	D	NA	G	G							+																											
Bugter-Maessen et al. (1996) [22]	P	Q*	1	G														+				+									+						
Bunting and van Walraven (2004) [16]	P	Q#	B	G																		+		+													
Burke (1995) [18]	D	NA	G	G																													+					
Carey and Garrett (1996) [9]	P	Q#	B	LBP												+									+	+											
Cherkin et al. (1994) [4]	P	Q#	2	LBP																			+														
DeKay and Asch (1998) [27]	D	NA	G	G																																	
Eccles et al. (2001) [7]	P	Q#	1	Rad																+																	
Elstein and Schwarz (2002) [24]	D	NA	G	G																																	
Espeland and Baerheim (2003) [5]	P	Q*	1	LBP			+					+		+					+																		
Espeland et al. (2001) [3]	P	B	2	LBP																																	
Feinstein (2002) [1]	D	NA	G	G		+		+							+							+														
Ferrier et al. (1991) [38]	P	Q*	1	G												+								+													
Freeborn et al. (1997) [35]	P	Q#	B	LBP																																	
Gallagher et al. (1997) [13]	P	Q#	1	MS										+											+												
Gosden et al. (1999) [34]	SR	Q#	B	G																																	
Grimshaw and Russell (1993) [21]	SR	Q#	B	G																																	
Kelly (1995) [19]	D	NA	G	Lab																+																	
Klein (2005) [23]	D	NA	G	G			+	+																													
Knottnerus (1991) [32]	D	NA	G	G		+	+																														
Knottnerus et al. (2002) [33]	D	NA	G	G		+				+	+																										
Kristiansen and Hjortdahl (1992) [6]	P	Q#	1	G														+																			
Little et al. (1998) [20]	P	Q#	1	LBP			+							+																							
Meyer et al. (2000) [26]	P	Q#	1	G																																	

Neilson et al. (2004) [25]	P	Q#	2	G	+																				+		
Pewsner et al. (2004) [29]	D	NA	G	G	+	+																					
Price (2000) [15]	D	NA	?	Lab	+	+		+																			
Ransohoff (2002) [30]	D	NA	G	G	+			+	+	+	+	+	+				+				+			+			
Roland and van Tulder (1998) [17]	D	NA	G	LBP						+				+					+								
Shye et al. (1998) [36]	P	Q*	B	LBP	+																+		+				+
Studdert et al. (2005) [37]	P	Q#	2	G													+										
van Der Weijden et al. (2003) [39]	P	Q#	1	G			+			+																	
Van wijk et al. (2001) [11]	P	Q#	1	Lab																					+		
Verstappen et al. (2004) [10]	P	Q#	1	G											+			+		+		+				+	
Verstappen et al. (2004) [8]	P	Q#	1	G														+							+		
Verstappen et al. (2003) [12]	P	Q#	1	G																					+		
Wong (1995) [14]	D	NA	G	Lab																						+	
Zaat and van Eijk (1992) [28]	P	Q#	1	Lab												+											

Analysis: NA = not applicable (discussion articles); Q* = qualitative; Q# = quantitative.
Setting: G = general; ? = unclear; 1 = primary care; 2 = secondary care.
Topic area: G = general; Lab = laboratory; rad = radiology.
[a] Study design: D = discussion article; SR = systematic review; P = primary study.

how they may influence test ordering for the diagnosis of MS and LBP.

4. Diagnostic factors

From a diagnostic perspective, tests are ordered to modify the pretest probability of disease [15,29,30,32,33,36], to rule in or rule out disease, or to refer for further evaluation [1,15,29,32,33]. In primary care tests are more commonly used to rule out a condition or to help the doctor to make a decision about referral, or provide further information on the patient to a secondary care specialist [5,20,32]. Another decision that primary care doctors commonly face is whether to order a test or to adopt a period of ‘‘watchful waiting,’’ asking a patient to return to see how his or her symptoms develop [39]. In contrast, in secondary care tests are more often used to reach a definite diagnosis [1,32]. Tests may also be used for triage to determine whether to refer a patient for further testing. Or, once initial tests have suggested the possibility of a particular condition, more costly or invasive tests may be ordered for confirmation. In such situations the possible differential diagnoses will dictate what tests may be offered.

Only a small proportion of people presenting with LBP have serious pathology and so the first step in the workup is to rule out conditions such as fracture (pretest probability 4%), spondylolisthesis (3%), herniated disk compression (1–3%), cancer (0.7%), or infection (0.01%) [40]. History and physical examination are used to identify ‘‘red flags,’’ which suggest the possibility of one of these conditions [40]. Laboratory tests, such as erythrocyte sedimentation rate or urinalysis, may be ordered to rule out cancer or infection. A positive result may indicate referral to secondary care or further testing, for example, radiography to confirm a diagnosis of cancer [41]. In patients with a nonspecific cause of LBP, imaging is not diagnostically indicated because it has poor specificity (anatomic abnormalities detected by imaging are often present in healthy individuals) and sensitivity (imaging often fails to identify causes of pain) [17,40]. Nevertheless, imaging is often undertaken in these patients [42]. The most commonly used test in the diagnosis of MS is magnetic resonance imaging (MRI) and to a lesser extent lumbar puncture and evoked potentials. From a diagnostic perspective, MRI is ordered to rule out serious treatable differential diagnoses, such as brain tumors, or to rule in MS.

5. Therapeutic and prognostic factors

Decisions regarding test ordering are interlinked with decisions regarding therapy and prognosis [33]. Although the focus of this article is on patients with undiagnosed symptoms, tests are often ordered to monitor the effects of interventions. Doctors may be wary of ordering

expensive and/or invasive tests if treatments are not available for the differential diagnoses. In contrast, tests may be ordered specifically to inform decisions on treatment such as whether to initiate therapy and what type to use, or to provide information on prognosis [1,30,31,33].

This is a difficult area in relation to MS: there is debate regarding the effectiveness of disease-modifying therapies for early MS [43]. Neurologists who believe that disease-modifying therapies are beneficial may order an MRI scan in an attempt to reach an earlier diagnosis so that patients can be prescribed these drugs sooner. An additional reason for ordering an MRI scan in patients with suspected MS is to provide a baseline for monitoring disease progression [44]. This may also help to provide information on prognosis after further testing. If back pain is because of an inflammatory disease such as ankylosing spondylitis, images of the sacroiliac joints and spine may provide a useful baseline from which to assess disease progression [45].

6. Patient-related factors

Factors related to the patient can influence test ordering in a variety of ways. Patient preference, either to undergo or not undergo testing or for one test over another, can influence test ordering [3,5,13,17,20,30]. The acceptability of the test to the patient, for example in terms of invasiveness or side effects, and the impact of the potential diagnosis on the patient may affect a doctor's decision to order a test [30]. The consequences of test results, in terms of false positive and false negative results as well as accurate results, should also be considered [17]. Doctors may decide to order a test to reassure the patient that he or she is not suffering from a more severe condition, for example ordering a scan in a patient with headache to reassure him or her that he or she does not have a brain tumor [1,20,30]. A number of studies have found that test ordering is related to patient demographics with, for example, older patients or female patients receiving a disproportionate number of tests [6,9,38].

Many patients with LBP or possible MS remain without a diagnostic label and continue to experience symptoms. A qualitative study reported that some patients with LBP ''feel that they have not been provided thorough work-ups and that additional tests, such as an x-ray or computed tomography scan, might allow precise ascertainment of the elusive physical cause of their suffering'' [46]. In such cases patients may request a test that is not diagnostically useful, and may be harmful, for example through exposure to gonadal radiation or impact on symptoms. A trial that randomized patients with nonspecific LBP to receive or not receive an x-ray found that patients in the x-ray group were more likely to report back pain after 3 months [42]. A possible explanation is the ''medicalizing'' effect of testing, which may increase patients' belief that they are unwell, leading to greater reporting of pain and limitation of activity. This study also found that, given the choice, over 80% of patients would have an x-ray. A review of three qualitative studies involving primary care doctors found that patient preference influenced their decision to order spinal radiography [5]. Patient requests for tests that are not indicated medically can be counteracted by providing information to explain this. A trial that randomized patients to receive immediate x-rays or an educational intervention found that after 3 weeks, 44% of those in the educational group believed that everyone with LBP should receive an x-ray compared to 73% in the x-ray group [47].

The effect that the test and its results may have on patients should be considered. Questionnaire studies of people with possible MS who underwent a full diagnostic workup found that patients who received a diagnosis of MS reported improved quality of life and decreased uncertainty, although patients felt less optimistic about their future health than before testing [48,49]. A positive test result may also have negative effects such as increased anxiety [48], increased insurance premiums, and workplace discrimination [43]. The consequence of inaccurate test results should also be considered. A false positive diagnosis of MS may result in unnecessary further tests and treatments, needless anxiety and psychological distress, and possible failure to identify a treatable condition [50]. A false negative result may mean that patients continue to experience unexplained symptoms, causing anxiety because they do not know what is wrong with them, or failure to get timely and appropriate treatment which may have adverse effects on their prognosis [48]. For other conditions, such as LBP, a false negative result may mean a delay in diagnosis and appropriate treatment which may have significant prognostic implications. For example, failure to diagnose and surgically correct a prolapsed vertebral disk may lead to urinary or fecal incontinence.

7. Doctor-related factors

Most factors identified as influencing test ordering are doctor-related factors. This is not surprising given that the decision to order a test in an individual patient ultimately rests with the individual doctor. Evidence suggests that doctors with more clinical experience, confidence in their clinical judgment [5,6,22], pride in their work [5], and who do not fear risk taking or uncertainty [1,5,16,22,28] tend to request fewer tests than those who dislike uncertainty and have less experience. Doctor demographics appear to influence test ordering. However, the results of two studies addressing this issue were contradictory: one found that younger doctors and male doctors tended to order fewer tests [6], whereas the other found that older doctors ordered fewer [26]. Both time pressures [5,30,36] and fear of litigation [5,27,30,37] have been found to result in increased testing. Individual doctors respond differently to patients' requests for unnecessary tests [5,13,17]. Involvement in

research, specifically in guideline development, was found to reduce the number of tests requested [10]. Another factor found to reduce test ordering was giving doctors feedback on how their test ordering rates compared to colleagues [8,10,16]. Cognitive biases influence the diagnostic process as a whole, and thus inevitably impact on test ordering [23,24]. For example, a recent consultation with a patient with a rare condition may predispose a doctor to suspect that diagnosis and request tests he or she may not otherwise have ordered. Additional factors found to influence test ordering were doctor specialty [4,9] and doctor working patterns [10,22].

A review of three qualitative studies of primary care doctors found that a variety of factors influenced their decisions to order radiography in patients with LBP: clinical skills, time pressure, risk of litigation, limiting conflict, ending difficult consultations, reducing their own anxiety, protecting professional pride, and pressure from other health care providers/organizations [5]. The findings of this review also suggest that differences in test ordering may be related to the geographical and/or cultural context in which doctors work. The three studies were conducted in the United States, the Netherlands and Norway, respectively, and showed that some of the barriers to general practioners' (GP) adherence to clinical guidelines regarding radiography ordering differed according to country. For example,

> ''GPs might order 'non-indicated' x-rays to buy time (USA), negotiate (all studies), or build a good relationship with the patient (USA, the Netherlands).''

An additional influence may be the doctor's perception of test accuracy. For example, a recent review has shown that MRI is not accurate either for ruling in or ruling out a diagnosis of MS when compared to the gold standard of long-term clinical follow-up [51]. Despite this, MRI is commonly requested in the evaluation of possible MS: a survey of Canadian neurologists found that 92% routinely order MRI scans in patients with suspected MS and 93% felt that MRI was ''very useful'' for the workup of these patients [52].

The way doctors respond to patients' requests for inappropriate testing and their ability to communicate the limitations of tests to patients may influence test ordering. A study of ''actor patients'' who presented with a complaint of fatigue and requested MRI scans to rule out MS, found that three of 39 doctors agreed to the MRI at the initial visit and eight said they might order MRI in the future, whereas the others explained that an MRI was not clinically indicated [13]. The type of specialty a doctor belongs to has been shown to affect test ordering for patients with LBP. Orthopedic surgeons were more likely to order a radiograph for patients with acute back pain than primary care doctors [9]. Another study reported that neurologists and neurosurgeons are twice as likely to request an imaging study for patients with acute non-radiating pain or chronic back pain than any other specialist [4]. This may be related to the different prior probabilities of disease or to differences in access to tests in these different settings.

8. Policy and organization-related factors

A number of influences on test ordering derive from local and national policy makers or the organization of the practice or hospital, and hence are outside the direct control of the individual doctor. For example in the United Kingdom, the limited availability of diagnostic tests in primary care is an important factor distinguishing it from secondary care, where diagnostic test availability is almost universal. A related factor that influences test ordering is the availability or ease of access to tests [5,22,36,38]. Payment by salaries is associated with lower test ordering than a fee for service approach where doctors are reimbursed for each test performed [6,26,30,34]. Type of primary care practice has been shown to influence test ordering in a number of studies, with doctors in solo or small practices tending to request more tests than those in group practices [9,10,22,26]. The introduction of a structured test ordering form has been found to reduce test ordering [7,10,22,26]. More broadly, policy makers issue guidelines regarding when tests should be used (http://www.nice.org.uk). In making these guidelines they consider not only the accuracy of the test but also its direct and indirect costs (http://www.nice.org.uk/pdf/GDM_Chapter8_0305.pdf). Direct costs include test and operator costs, whereas indirect costs are broader including consequences of test results in terms of treatment, implications of inaccurate test results, and longer-term costs in terms of patient outcome. A number of studies have found that clinical guidelines and policy recommendations affect doctors' test ordering, generally by reducing inappropriate test ordering [8,11,12,18,19,21,25,35].

A review of three qualitative studies of primary care doctors' reasons for ordering radiography in patients with LBP found that policy and organizational factors, such as ease of access to testing, influenced their decisions [5]. Because health care policy and organization differs in different countries, doctors' geographical location and/or cultural context is also likely to play a role. As discussed above, the aforementioned review supported this hypothesis. For example, the three studies showed that ''easy access to actual x-ray services (Norway), and difficult access to physiotherapy (United States) or computed tomography (Norway)'' influenced test ordering practices. A study that looked at factors that affected doctors' decisions to order lumbar spine x-rays in patients with LBP found that ambiguity about the internal referral process was an important factor [36]. There is evidence that interventions at the organizational level can reduce inappropriate test ordering. Randomized clinical trials have found that attaching educational reminder messages to radiography improved referral for imaging tests [7] and that providing specific

guidelines led to improved test ordering compared to non-specific guidelines [53].

Clinical guidelines regarding the diagnosis of MS and LBP have been developed for the National Health Service (NHS) in England and Wales. The National Institute for Health and Clinical Excellence clinical guidelines for MS recommend that if a diagnosis cannot be made on clinical grounds alone the McDonald criteria, which incorporate MRI, should be used to find evidence that supports the diagnosis (http://www.nice.org.uk/pdf/CG008guidance.pdf).

Similarly, in the United States, the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology and the Practice Committee of the Child Neurology Society have issued guidance on the use of MRI in the diagnosis of MS [54]. NHS guidance on the management of LBP in primary care recommends that diagnostic imaging should only be considered if there are clinical indications of serious pathology (http://www.prodigy.nhs.uk/guidance.asp?gt=back%20pain%20-%20lower). Policy makers and guideline panels appreciate the need to limit unnecessary diagnostic testing. It is important to understand that the extent and strength of their recommendations are often modified by other factors aside from the accuracy of the diagnostic test, particularly cost, availability, and potential hazard.

9. Discussion

This review identified five key interrelated factor groupings that influence a doctor's decision to order a test for a particular patient: diagnostic factors, therapeutic and prognostic factors, patient-related factors, doctor-related factors, and policy and organization-related factors. Because test accuracy is only one of a number of considerations, attempts to influence test ordering practice are unlikely to succeed unless they account for the wider context implied by these factors.

Previous review articles have tended to focus only on particular aspects of test ordering: for example, the effects of cognitive biases on doctors' diagnostic reasoning [23], methods of doctor payment [34], and clinical guidelines [21]. Other discussion articles have focused exclusively on diagnostic factors [29,32]. While a number of articles have covered more than one aspect of test ordering, to our knowledge, this review is the first to provide an evidence-based overview of the different factors that may influence test ordering.

Although this review was not based on exhaustive searches of the literature, we used systematic methods throughout the review process. We used explicit inclusion criteria and a structured data extraction form and all aspects of the review process were double-checked. The included studies used a variety of different terms to discuss test ordering and many studies had multiple aims; investigating influences on test ordering was not always the primary objective. Identifying such studies, in a situation in which there are no clear search terms, is very difficult. Given that the objective of this review was to identify reasons for test ordering rather than to quantify the evidence for the effects of different factors on test ordering, we felt that it was not necessary to include all relevant studies. We used the standard qualitative technique of searching until saturation was reached and no new themes emerged. Although this gives us confidence that the key factors that may influence test ordering have been included in this review, its main limitation is the lack of a fully systematic search strategy, aiming to include every relevant study. Although the review includes studies conducted in a number of countries (e.g., United Kingdom, United States, The Netherlands, and Norway), the inclusion only of English language articles is a specific limitation. The variety of study designs and the very broad range of topics covered also made it difficult to review the located studies, and to synthesize data from them. We overcame this problem by identifying the key factors that influence test ordering and then using two specific examples, LBP and MS, to summarize how these factors may operate in practice.

This review has implications for three key areas: test accuracy research, test ordering research, and test ordering behavior. Test accuracy research, which focuses on determining the diagnostic accuracy of a test, should also consider the broader context in which test ordering takes place in practice. This includes both the aim of testing (the context in which the test will be used, e.g., triage, add-on or replacement) as well as the social and psychological influences identified by this review. This applies to both primary studies and systematic reviews. There are two ways in which this could be done. A simple approach is to acknowledge the additional factors that may influence test ordering when making recommendations for practice. For example, in a recent review on the accuracy of MRI for the diagnosis of MS, we included a discussion of the non-diagnostic reasons for ordering an MRI scan [51]. A more robust method would be to incorporate both clinical expertise and qualitative findings on the broader clinical picture into test accuracy research. This could be done either by reviewing existing research or by carrying out additional primary research.

Test ordering research considers the different factors that may influence decisions to order tests in practice. This review provides an overview of this area of research and forms the basis for future research, of which a key aim would be to quantify the effects of different influences. This could be done by carrying out a systematic review of all primary studies that have assessed influences on test ordering. The aim of the review would be to quantify the volume of evidence for the different influences on test ordering, and also to determine the magnitude of each of these influences. Such a review should carry out exhaustive literature searches and include an assessment of the quality of the included studies. Future research should also consider how

the very different factors influencing test ordering may influence the overall pattern of test ordering, not just by their individual influences but also by their interactions. A doctor may choose to order a test with relatively poor diagnostic performance for a range of interacting reasons. Considering these influences in isolation would miss understanding the full picture. For example, electrocardiograms (ECG) are commonly ordered in patients presenting with a variety of symptoms (e.g., chest pain, shortness of breath, dizzy spells, palpitations) even though the test's value in diagnosing these potential underlying conditions (e.g., myocardial infarction, angina, heart failure, intermittent heart beat irregularity) is limited because of poor specificity. Despite this limitation, ECG testing has several characteristics (ease of access, relative low cost, absence of adverse effects ease of interpretation, and instant availability of results) that may all contribute to a doctor's decision to order this test.

The ultimate aim of test accuracy and test ordering research is to improve the use of diagnostic tests in practice, that is, test ordering behavior. This review suggests a number of areas where interventions to improve test ordering behavior may be considered: for example, the introduction of structured test ordering forms or feeding back individual doctors' test ordering patterns. In addition, as this review highlights, decisions regarding test ordering regularly involve more than a mechanistic consideration of a test's properties and clinical appropriateness (e.g., a doctor's belief that a test is inappropriate may have to be weighed against a patient's need for reassurance). Sometimes, such subjective factors may override the scientific evidence regarding diagnostic accuracy, as a result of the doctor's considered judgment of the case as a whole.

In conclusion, the decision to order a test rests with the doctor in consultation with the patient. We have shown that a wide range of factors will affect this decision. Ignoring these factors and concentrating only on test accuracy risks missing the realities of diagnosis in clinical practice.

References

[1] Feinstein AR. Misguided efforts and future challenges for research on "diagnostic tests". J Epidemiol Community Health 2002;56:330–2.

[2] Dixon-Woods M, Agarwal S, Jones D, Young B, Sutton A. Synthesising qualitative and quantitative evidence: a review of possible methods. J Health Serv Res Policy 2005;10:45–53.

[3] Espeland A, Baerheim A, Albrektsen G, Korsbrekke K, Larsen JL. Patients' views on importance and usefulness of plain radiography for low back pain. Spine 2001;26:1356–63.

[4] Cherkin DC, Deyo RA, Wheeler K, Ciol MA. Physician variation in diagnostic testing for low back pain. Who you see is what you get. Arthritis Rheum 1994;37:15–22.

[5] Espeland A, Baerheim A. Factors affecting general practitioners' decisions about plain radiography for back pain: implications for classification of guideline barriers – a qualitative study. BMC Health Serv Res 2003;3:8.

[6] Kristiansen IS, Hjortdahl P. The general practitioner and laboratory utilization: why does it vary? Fam Pract 1992;9:22–7.

[7] Eccles M, Steen N, Grimshaw J, Thomas L, McNamee P, Soutter J, et al. Effect of audit and feedback, and reminder messages on primary-care radiology referrals: a randomised trial. Lancet 2001;357:1406–9.

[8] Verstappen WHJM, van Merode F, Grimshaw J, Dubois WI, Grol RPTM, van Der Weijden T. Comparing cost effects of two quality strategies to improve test ordering in primary care: a randomized trial. Int J Qual Health Care 2004;16:391–8.

[9] Carey TS, Garrett J. Patterns of ordering diagnostic tests for patients with acute low back pain. The North Carolina Back Pain Project. Ann Intern Med 1996;125:807–14.

[10] Verstappen WH, ter Riet G, Dubois WI, Winkens R, Grol RP, van Der Weijden T. Variation in test ordering behaviour of GPs: professional or context-related factors? Fam Pract 2004;21:387–95.

[11] van Wijk MAM, van der Lei J, Mosseveld M, Bohnen AM, van Bemmel JH. Assessment of decision support for blood test ordering in primary care: a randomized trial. Ann Intern Med 2001;134:274–81.

[12] Verstappen WHJM, van Der Weijden T, Sijbrandij J, Smeele I, Hermsen J, Grimshaw J, et al. Effect of a practice-based strategy on test ordering performance of primary care physicians: a randomized trial. JAMA 2003;289:2407–12.

[13] Gallagher TH, Lo B, Chesney M, Christensen K. How do physicians respond to patient's requests for costly, unindicated services? J Gen Intern Med 1997;12:663–8.

[14] Wong ET. Improving laboratory testing: can we get physicians to focus on outcome? Clin Chem 1995;41:1241–7.

[15] Price CP. Evidence-based laboratory medicine: supporting decision-making. Clin Chem 2000;46:1041–50.

[16] Bunting PS, van Walraven C. Effect of a controlled feedback intervention on laboratory test ordering by community physicians. Clin Chem 2004;50:321–6.

[17] van Roland M, van Tulder M. Should radiologists change the way they report plain radiography of the spine? Lancet 1998;352:229–30.

[18] Burke MD. Clinical laboratory consultation. Clin Chem 1995;41:1237–40.

[19] Kelly JT. Role of clinical practice guidelines and clinical profiling in facilitating optimal laboratory use. Clin Chem 1995;41:1234–6.

[20] Little P, Cantrell T, Roberts L, Chapman J, Langridge J, Pickering R. Why do GPs perform investigations?: the medical and social agendas in arranging back X-rays. Fam Pract 1998;15:264–5.

[21] Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. Lancet 1993;342:1317–22.

[22] Bugter-Maessen AM, Winkens RA, Grol RP, Knottnerus JA, Kester AD, Beusmans GH, et al. Factors predicting differences among general practitioners in test ordering behaviour and in the response to feedback on test requests. Fam Pract 1996;13:254–8.

[23] Klein JG. Five pitfalls in decisions about diagnosis and prescribing. BMJ 2005;330:781–3.

[24] Elstein AS, Schwarz A. Evidence base of clinical diagnosis: clinical problem solving and diagnostic decision making: selective review of the cognitive literature. BMJ 2002;324:729–32.

[25] Neilson EG, Johnson KB, Rosenbloom ST, Dupont WD, Talbert D, Giuse DA, et al. The impact of peer management on test-ordering behavior. Ann Intern Med 2004;141:196–204.

[26] Meyer CM, Ladenson PW, Scharfstein JA, Danese MD, Powe NR. Evaluation of common problems in primary care: effects of physician, practice, and financial characteristics. Am J Manag Care 2000;6:457–69.

[27] DeKay ML, Asch DA. Is the defensive use of diagnostic tests good for patients, or bad? Med Decis Making 1998;18:19–28.

[28] Zaat JO, van Eijk JT. General practitioners' uncertainty, risk preference, and use of laboratory tests. Med Care 1992;30:846–54.

[29] Pewsner D, Battaglia M, Minder C, Marx A, Bucher HC, Egger M. Ruling a diagnosis in or out with "SpPIn" and "SnNOut": a note of caution. BMJ 2004;329:209–13.

[30] Ransohoff DF. Challenges and opportunities in evaluating diagnostic tests. J Clin Epidemiol 2002;55:1178–82.

[31] Asch DA, Patton JP, Hershey JC. Knowing for the sake of knowing: the value of prognostic information. Med Decis Making 1990;10: 47−57.
[32] Knottnerus JA. Medical decision making by general practitioners and specialists. Fam Pract 1991;8:305−7.
[33] Knottnerus JA, van WC, Muris JW. Evaluation of diagnostic procedures. BMJ 2002;324:477−80.
[34] Gosden T, Pedersen L, Torgerson D. How should we pay doctors? A systematic review of salary payments and their effect on doctor behaviour. Q J Med 1999;92:47−55.
[35] Freeborn DK, Shye D, Mullooly JP, Eraker S, Romeo J. Primary care physicians' use of lumbar spine imaging tests: effects of guidelines and practice pattern feedback. J Gen Intern Med 1997;12: 619−25.
[36] Shye D, Freeborn DK, Romeo J, Eraker S. Understanding physicians' imaging test use in low back pain care: the role of focus groups. Int J Qual Health Care 1998;10:83−91.
[37] Studdert DM, Mello MM, Sage WM, DesRoches CM, Peugh J, Zapert K, et al. Defensive medicine among high-risk specialist physicians in a volatile malpractice environment. JAMA 2005;293: 2609−17.
[38] Ferrier BM, Woodward CA, Cohen M, Goldsmith CH. Laboratory tests: which physicians order more? Can Fam Phys 1991;37:349−52.
[39] van Der Weijden T, van Velsen M, Dinant GJ, van Hasselt CM, Grol R. Unexplained complaints in general practice: prevalence, patients' expectations, and professionals' test-ordering behavior. Med Decis Making 2003;23:226−31.
[40] Atlas SJ, Deyo RA. Evaluating and managing acute low back pain in the primary care setting. J Gen Intern Med 2001;16:120−31.
[41] Jarvik JG, Hollingworth W, Martin B, Emerson SS, Gray DT, Overman S, et al. Rapid magnetic resonance imaging vs radiographs for patients with low back pain: a randomized controlled trial. JAMA 2003;289:2810−8.
[42] Kendrick D, Fielding K, Bentley E, Kerslake R, Miller P, Pringle M. Radiography of the lumbar spine in primary care patients with low back pain: randomised controlled trial. BMJ 2001;322:400−5.
[43] O'Connor P. Canadian Multiple Sclerosis Working Group. Key issues in the diagnosis and treatment of multiple sclerosis. An overview. Neurology 2002;59(6 Suppl 3):S1−S33.
[44] Stevenson VL, Miller DH. Magnetic resonance imaging in the monitoring of disease progression in multiple sclerosis. Mult Scler 1999;5:268−72.
[45] Braun J, Baraliakos X, Golder W, Brandt J, Rudwaleit M, Listing J, et al. Magnetic resonance imaging examinations of the spine in patients with ankylosing spondylitis, before and after successful therapy with infliximab: evaluation of a new scoring system. Arthritis Rheum 2003;48:1126−36.
[46] Borkan J, Reis S, Hermoni D, Biderman A. Talking about the pain: a patient-centered study of low back pain in primary care. Soc Sci Med 1995;40:977−88.
[47] Deyo RA, Diehl AK, Rosenthal M. Reducing roentgenography use. Can patient expectations be altered? Arch Intern Med 1987;147: 141−5.
[48] Mushlin AI, Mooney C, Grow V, Phelps CE. The value of diagnostic information to patients with suspected multiple sclerosis. Rochester−Toronto MRI Study Group. Arch Neurol 1994;51:67−72.
[49] O'Connor P, Detsky AS, Tansey C, Kucharczyk W. Effect of diagnostic testing for multiple sclerosis on patient health perceptions. Rochester−Toronto MRI Study Group. Arch Neurol 1994;51:46−51.
[50] Rudick RA, Schiffer RB, Schwetz KM, Herndon RM. Multiple sclerosis. The problem of incorrect diagnosis. Arch Neurol 1986;43: 578−83.
[51] Whiting P, Harbord J, Main C, Deeks J, Filippini G, Egger M, et al. The accuracy of MRI for the diagnosis of MS: a systematic review. BMJ 2006;332:875−84.
[52] O'Connor P, Lee L. Access to multiple sclerosis diagnosis for Canadian neurologists. Can J Neurol Sci 1999;26:115−8.
[53] Shekelle PG, Kravitz RL, Beart J, Marger M, Wang M, Lee M. Are nonspecific practice guidelines potentially harmful? A randomized comparison of the effect of nonspecific versus specific guidelines on physician decision making. Health Serv Res 2000;34:1429−48.
[54] Frohman EM, Goodin DS, Calabresi PA, Corboy JR, Coyle PK, Filippi M, et al. Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology. The utility of MRI in suspected MS: report of the therapeutics and technology assessment subcommittee of the American Academy of Neurology. Neurology 2003;61:602−11.